ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Complexity 21 (2005) 479–486
Journal of COMPLEXITY
www.elsevier.com/locate/jco

Generalized Budan–Fourier theorem and virtual roots

Michel Coste[a], Tomás Lajous-Loaeza[a,b], Henri Lombardi[c,*], Marie-Françoise Roy[d]

[a] IRMAR (URA CNRS 305), Université de Rennes, Campus de Beaulieu 35042 Rennes Cedex, France
[b] UBS Investment Research, Campos El'ıseos 345-19, Col Polanco, 11560 México, DF, México
[c] Laboratoire de Mathématiques de Besançon, (UMR CNRS 6623), Université de Franche-Comté, 25030 Besancon Cedex, France
[d] IRMAR (URA CNRS 305), Université de Rennes, Campus de Beaulieu 35042 Rennes Cedex, France

Received 31 May 2004; accepted 28 November 2004
Available online 2 March 2005

Abstract

In this Note we give a proof of a generalized version of the classical Budan–Fourier theorem, interpreting sign variations in the derivatives in terms of virtual roots.

Keywords: Budan–Fourier rule; Real roots; Virtual roots; Fewnomials

0. Introduction

The notion of virtual root was introduced in [5] in the case of polynomials. The virtual roots provide d continuous "root fonctions" on the space all real polynomials of a given degree d, with an interlacing property linking virtual roots of P and virtual roots of P'. From a computational point of view, there is no need to know the coefficients with infinite precision in order to compute the virtual roots with finite precision: the discontinuity phenomenon of real roots vanishing in the complex plane "disappears". Another nice fact is that all real roots of P are virtual roots and all virtual roots are real roots of P or of one of its derivatives.

* Corresponding author. Fax: +33 3 81 666526.
E-mail addresses: michel.coste@univ-rennes1.fr (M. Coste), tomas.lajous@ubs.com (T. Lajous-Loaeza), henri.lombardi@univ-fcomte.fr (H. Lombardi), marie-francoise.roy@univ-rennes1.fr (M.-F. Roy).

doi:10.1016/j.jco.2004.11.003

Theorem 2 of this note shows that the Budan–Fourier count always gives the number of virtual roots (with multiplicities) on an interval $(a, b]$. So the quantity by which the Budan–Fourier count exceeds the number of actual roots is explained by the presence of extra virtual roots.

The Budan–Fourier count of virtual roots is a useful addition to [5]. It gives a way to obtain approximations of the virtual roots, by dichotomy, merely by evaluation of signs of derivatives.

We formulate and prove the results in a more general framework than the one of polynomials, by introducing the notions of degree and successive derivatives with respect to a sequence $\mathbf{f}$ of positive functions. The motivating example for introducing this enlarged framework is given by "fewnomials": one takes into account the number of monomials (possibly with real exponents), and not the actual degree. The framework can also be applied to exponential polynomials, for instance.

The nice properties of virtual roots of polynomials remain essentially the same for $\mathbf{f}$-virtual roots. We have for instance the interlacing property (Proposition 2). The main result is that the Budan–Fourier count (with $\mathbf{f}$-derivatives) gives the exact number of $\mathbf{f}$-virtual roots on an interval, and, hence, an upper bound on the number of actual roots (Theorem 1).

1. Generalized Budan–Fourier theorem

The *number of sign changes*, $\mathrm{V}(a)$, in a sequence, $a = a_0, \ldots, a_p$, of elements in $\mathbb{R} \setminus \{0\}$ is defined by induction on p by

$$\mathrm{V}(a_0) = 0,$$
$$\mathrm{V}(a_0, \ldots, a_p) = \begin{cases} \mathrm{V}(a_1, \ldots, a_p) + 1 & \text{if } a_0 a_1 < 0, \\ \mathrm{V}(a_1, \ldots, a_p) & \text{if } a_0 a_1 > 0. \end{cases}$$

This definition extends to any finite sequence a of elements in $\mathbb{R}$ by considering the finite sequence b obtained by dropping the zeros in a and defining $\mathrm{V}(a) = \mathrm{V}(b)$, stipulating that V of the empty sequence is 0.

In the following we fix an interval (α, β), with $\alpha \in \mathbb{R} \cup \{-\infty\}$, $\beta \in \mathbb{R} \cup \{+\infty\}$ and we denote by $\mathcal{C}^\infty(\alpha, \beta)$ the ring of infinitely differentiable functions on (α, β). Let $\mathcal{G} = g_0, g_1, \ldots, g_N$ be a sequence of elements of $\mathcal{C}^\infty(\alpha, \beta)$ and let a be an element of (α, β). The *number of sign changes* of $\mathcal{G}$ at a, denoted by $\mathrm{V}(\mathcal{G}; a)$, is $\mathrm{V}(g_0(a), \ldots, g_N(a))$. Given a and b in (α, β) we write $\mathrm{V}(\mathcal{G}; a, b)$ for $\mathrm{V}(\mathcal{G}; a) - \mathrm{V}(\mathcal{G}; b)$.

Let $\mathbf{f} = (f_0, f_1, \ldots)$ be an infinite sequence of positive functions in $\mathcal{C}^\infty(\alpha, \beta)$. For $i \in \mathbb{N}$, define the operator D_i on $\mathcal{C}^\infty(\alpha, \beta)$ by $\mathrm{D}_i g = (g/f_i)'$. By definition, a function $g \in \mathcal{C}^\infty(\alpha, \beta)$ is *of* $\mathbf{f}$*-degree* d if and only if

$$\mathrm{D}_0 \mathrm{D}_1 \cdots \mathrm{D}_d g = 0 \quad \text{and} \quad \mathrm{D}_1 \cdots \mathrm{D}_d g \neq 0 \quad (g \neq 0 \text{ in case } d = 0).$$

The $\mathbf{f}$-degree of a function might not be uniquely defined; we implicitly assume in this case that the $\mathbf{f}$-degree is given with the function.

Let g be a function of $\mathbf{f}$-degree d. The *list* $\mathrm{Der}_{\mathbf{f}}(g)$ *of* $\mathbf{f}$*-derivatives of* g is by definition the list $g_0, \ldots, g_d$, where $g_d = g$ and $g_{i-1} = \mathrm{D}_i g_i$ for $i = d, \ldots, 1$. Note that g_i is of $\mathbf{f}$-degree i.

The $\mathbf{f}$-*multiplicity of a root a of g* is the index m such that $g_d(a) = \cdots = g_{d-m+1}(a) = 0,\ g_{d-m}(a) \neq 0$. In particular, if a is not a root of g, its $\mathbf{f}$-multiplicity is 0.

Note that g changes sign at a if and only if the $\mathbf{f}$-multiplicity of a is odd. This can be checked by induction on the $\mathbf{f}$-multiplicity, using the variation of g/f_d.

Example 1. Take $(\alpha, \beta) = \mathbb{R}$ and $\mathbf{f} = (1, 1, \ldots)$. The D_i are the usual derivatives, and a function of $\mathbf{f}$-degree d is, of course, a polynomial of degree d. We recover the usual notion of list of derivatives of g and the usual notion of multiplicity.

The following example is the motivating example for the introduction of the $\mathbf{f}$-degree and of the $\mathbf{f}$-derivatives. Here the $\mathbf{f}$-degree is the number of monomials (possibly with real exponents) minus 1, and the successive $\mathbf{f}$-derivative decrease the number of monomials by one.

Example 2. We start with

$$g(x) = a_d x^{r_d} + a_{d-1} x^{r_{d-1}} + \cdots + a_0 x^{r_0},$$

where $a_d \neq 0$ and $(r_0, \ldots, r_d)$ is an increasing sequence of real numbers. The function g is in $\mathcal{C}^\infty((0, +\infty))$. Take

$$\mathbf{f} = (x^{r_d - r_{d-1} - 1}, x^{r_{d-1} - r_{d-2} - 1}, \ldots, x^{r_1 - r_0 - 1}, x^{r_0}, 1, 1, \ldots).$$

Then g is of $\mathbf{f}$-degree d and its $\mathbf{f}$-derivatives are

$$\begin{aligned}
g_{d-1}(x) &= a_d(r_d - r_0)x^{r_d - r_0 - 1} + a_{d-1}(r_{d-1} - r_0)x^{r_{d-1} - r_0 - 1} + \cdots \\
&\quad + a_1(r_1 - r_0)x^{r_1 - r_0 - 1} \\
&\ \ \vdots \\
g_1(x) &= a_d(r_d - r_0)(r_d - r_1)\cdots(r_d - r_{d-2})x^{r_d - r_{d-2} - 1} \\
&\quad + a_{d-1}(r_{d-1} - r_0)(r_{d-1} - r_1)\cdots(r_{d-1} - r_{d-2})x^{r_{d-1} - r_{d-2} - 1} \\
g_0(x) &= a_d(r_d - r_0)(r_d - r_1)\cdots(r_d - r_{d-1})x^{r_d - r_{d-1} - 1}.
\end{aligned}$$

In particular we have $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); x) = \mathrm{V}(a_0, a_1, \ldots, a_d)$ for every $x > 0$ sufficiently small.

Example 3. Take

$$\mathbf{f} = (1, 1, 1, 1, 1, e^{ax}, 1, 1, 1, \ldots).$$

where a is a nonzero real number. Then a function

$$g(x) = P(x) + Q(x)e^{ax} \in \mathcal{C}^\infty(\mathbb{R}),$$

where P, $Q \in \mathbb{R}[x]$ and $\deg(Q) = 5$ is of $\mathbf{f}$-degree $6 + \deg(P)$.

This example may be easily generalized in many ways.

We are going to prove the following result:

Theorem 1 (*Generalized Budan–Fourier theorem*). *Let* $g \in \mathcal{C}^{\infty}(\alpha, \beta)$ *be of* **f**-*degree* d. *Denote by* $n(g; (a, b])$ *the number of roots of* g *in* $(a, b]$ *counted with* **f**-*multiplicities. Then*:

- $n(g; (a, b])$ *is not greater than* $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b)$ (*which is obviously* $\leqslant d$).
- $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b) - n(g; (a, b])$ *is even*.

The classical Budan–Fourier theorem [3,4] generalizes the Descartes rule [2] and is a particular case, corresponding to the situation of Example 1. Generalized Budan–Fourier theorem in the particular case of Example 2 was already known to Sturm [6].

A natural question is to interpret the difference $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b) - n(g; (a, b])$. This is done through virtual roots. The consideration of virtual roots will also provide a proof of Theorem 1.

2. Definition and properties of generalized virtual roots

The notion of virtual root was introduced in [5] in the case of polynomials.

We generalize the definition of virtual roots and precise the notion of virtual multiplicity in our context of **f**-derivatives. We first define the notion of **f**-virtual multiplicity, proceeding by induction on the **f**-degree.

Definition 1. Let g be a function of **f**-degree 0. For every $a \in (\alpha, \beta)$, the **f**-*virtual multiplicity of a as a root of g* is 0.

Let g be a function of **f**-degree $d > 0$. Let a be a point in (α, β), and let v be the **f**-virtual multiplicity of a as a root of $\mathrm{D}_d g$. Then the **f**-*virtual multiplicity of a as a root of g* is

- $v + 1$ if $g(x)\mathrm{D}_d g(x)$ is negative for $x < a$ close to a and positive for $x > a$ close to a (this means that $g(x)/f_d(x)$ goes away from 0 as x leaves a in both directions, i.e., $|g/f_d|$ has a local minimum at a),
- v if $g(x)\mathrm{D}_d g(x)$ does not change sign at a (this means that $g(x)/f_d(x)$ goes closer to 0 as x leaves a in one direction, and away from 0 as x leaves a in the other direction, i.e., $|g/f_d|$ is monotone near a),
- $v - 1$ if $g(x)\mathrm{D}_d g(x)$ is positive for $x < a$ close to a and negative for $x > a$ close to a (this means that $g(x)/f_d(x)$ goes closer to 0 as x leaves a in both directions, i.e., $|g/f_d|$ has a local maximum at a).

Note that if a is a root of g, the **f**-virtual multiplicity of a as a root of g is necessarily $v + 1$.

Example 4. Let us take an example in the case of polynomials (as in Example 1). Let P be a polynomial of degree 2.

- If P has two distinct real roots, they have virtual multiplicity 1 and the virtual multiplicity is zero elsewhere.

- Otherwise, the root of P' has virtual multiplicity 2 as root of P (even if it is not a root of P), and the virtual multiplicity is zero elsewhere.

It is not obvious from the definition that the $\mathbf{f}$-virtual multiplicity is everywhere nonnegative. This follows from the next proposition.

Proposition 1. *Let g be of $\mathbf{f}$-degree d. For every $a \in (\alpha, \beta)$, the $\mathbf{f}$-virtual multiplicity of a as a root of g is not less than its $\mathbf{f}$-multiplicity, and the difference is even.*

Proof. The proof is by induction on d. The claim obviously holds if $d = 0$. We assume now that $d > 0$ and the claim holds for $d - 1$; in particular, it holds for D_{dg}. We denote by v the $\mathbf{f}$-virtual multiplicity of a as a root of $\mathrm{D}_d g$ and by μ its $\mathbf{f}$-multiplicity; we have $v \geqslant \mu$ and $v - \mu$ even.

- If a is a root of g, then its $\mathbf{f}$-virtual multiplicity as a root of g is $v + 1$ and its $\mathbf{f}$-multiplicity is $\mu + 1$. The claim holds in this case.
- If a is not a root of g and v and μ are even, then $\mathrm{D}_d g$ does not change sign at a and g/f_d is monotonic in a neighborhood of a. Hence, the $\mathbf{f}$-virtual multiplicity of a as a root of g is v and its $\mathbf{f}$-multiplicity is 0. The claim holds in this case.
- If a is not a root of g and v and μ are odd, then $\mathrm{D}_d g$ changes sign at a. It follows that the $\mathbf{f}$-virtual multiplicity of a as a root of g is $v + 1$ or $v - 1$, while its $\mathbf{f}$-multiplicity is 0. The claim also holds in this case. □

We can now define the $\mathbf{f}$-virtual roots.

Definition 2. Let g be a function of $\mathbf{f}$-degree d on (α, β). A real a in (α, β) is *a* $\mathbf{f}$*-virtual root of* g if and only if it has a positive $\mathbf{f}$-virtual multiplicity as root of g.

Note that every root of g is a $\mathbf{f}$-virtual root, and that every $\mathbf{f}$-virtual root of g is an actual root of g or of one of its $\mathbf{f}$-derivatives.

When g is a polynomial of degree d having d distinct real roots, Rolle's theorem implies that these roots are interlaced with the roots of the derivative g'. This interlacing property holds in all cases for virtual roots [5]. We make precise this interlacing property in the context of $\mathbf{f}$-derivatives.

Proposition 2. *Let g be a function of $\mathbf{f}$-degree $d > 0$ on (α, β), and assume that $\mathrm{D}_d g$ has at least one $\mathbf{f}$-virtual root. Order the $\mathbf{f}$-virtual roots of $\mathrm{D}_d g$ as*

$$b_1 \leqslant b_2 \leqslant \cdots \leqslant b_n,$$

with repetitions according to $\mathbf{f}$-virtual multiplicities.

- *Assume $g(x)/f_d(x)$ goes away from 0 as x leaves b_1 on the left side, or g has a root in (α, b_1). Assume also that $g(x)/f_d(x)$ goes away from 0 as x leaves b_n on the right side, or g has a root in (b_n, β). Then there are $n + 1$ $\mathbf{f}$-virtual roots $a_1 \leqslant \cdots \leqslant a_{n+1}$ of*

g counted with $\mathbf{f}$*-virtual multiplicity, and they are interlaced with the* $\mathbf{f}$*-virtual roots of* $\mathrm{D}_d g$ *as follows*:

$$a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \cdots \leqslant a_n \leqslant b_n \leqslant a_{n+1}.$$

- *If $g(x)/f_d(x)$ goes closer to* 0 *as x leaves b_1 on the left side and g has no root in* (α, b_1), *delete a_1.*
- *If $g(x)/f_d(x)$ goes closer to* 0 *as x leaves b_n on the right side and g has no root in* (b_n, β), *delete a_{n+1}.*

Proof. First, recall that a $\mathbf{f}$-virtual root c of $\mathrm{D}_d g$ of $\mathbf{f}$-virtual multiplicity v contributes for $v-1$ to the $\mathbf{f}$-virtual multiplicity as a root of g, plus one for each side of c along which $g(x)/f_d(x)$ goes away from 0 as x leaves c.

Now consider distinct consecutive $\mathbf{f}$-virtual roots c and d of $\mathrm{D}_d g$. The function g/f_d is monotonic on (c, d). Hence, either there is a root of g in (c, d), or $g(x)/f_d(x)$ goes away from 0 as x leaves c on the right side, or $g(x)/f_d(x)$ goes away from 0 as x leaves d on the left side, and these three possibilities are exclusive.

The interlacing property follows from these considerations. $\square$

Note that we have a property of alternation of signs: There is $\sigma \in \{-1, 1\}$ such that, for every $x \in (\alpha, \beta)$ such that $a_i < x < a_{i+1}$, the sign of $g(x)$ is $(-1)^i\sigma$ (it is understood that $0 \leqslant i \leqslant d$ and that the corresponding inequality is deleted when a_i or a_{i+1} does not exist). This follows from the relation between the change of sign of $g(x)$ and the parity of the $\mathbf{f}$-multiplicity of roots, and from the fact that the difference between $\mathbf{f}$-virtual multiplicity and $\mathbf{f}$-multiplicity is even.

3. Virtual roots and sign variations in derivatives

We denote by $v(g; (a, b])$ the number of $\mathbf{f}$-virtual roots of g in $(a, b]$, counted with $\mathbf{f}$-virtual multiplicities. By Proposition 1, we have $v(g; (a, b]) \geqslant n(g; (a, b])$, and the difference is even. The section is devoted to the proof of the following result.

Theorem 2. *Let g be a function of* $\mathbf{f}$*-degree d, and let $a < b$ be real numbers in* (α, β). *Then*:

$$v(g; (a, b]) = \mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b).$$

In particular, the number of $\mathbf{f}$*-virtual roots of g counted with* $\mathbf{f}$*-virtual multiplicities is at most d.*

The generalized Budan–Fourier theorem (Theorem 1) follows immediately from Proposition 1 and Theorem 2. The following lemmas are the key to the proof of Theorem 2. In these lemmas, g is a function of $\mathbf{f}$-degree d, and we denote by $g_d = g,\ g_{d-1}, \ldots, g_0$ the $\mathbf{f}$-derivatives of g.

Lemma 1. *Let $a \in (\alpha, \beta)$. There is $\varepsilon > 0$ such that, for every $x \in (a, a+\varepsilon)$,*

$$\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, x) = 0,$$
$$v(g; (a, x]) = 0.$$

Proof. The first part of the assertion is the only one which deserves a proof. If $g_i(a) \neq 0$ for $i = 0, \ldots, d$, the claim is obvious. If $g_i(a) \neq 0$ and $g_{i+1}(a) = \cdots = g_{i+\ell}(a) = 0$, then $g_{i+1}(x), \ldots, g_{i+\ell}(x)$ all have the same sign as $g_i(x)$ for $x > a$ close to a; hence the number of sign changes remains the same, and the claim is proved. □

Lemma 2. *Let $c \in (\alpha, \beta)$ have $\mathbf{f}$-virtual multiplicity w as a root of g. Choose a and b such that $\alpha < a < c < b < \beta$ and no g_i, $0 \leqslant i \leqslant d$, has a root in $[a, c) \cup (c, b]$. Then*

$$\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b) = w.$$

Proof. The proof of the claim is by induction on the $\mathbf{f}$-degree d of g. The claim obviously holds if $d = 0$. Now assume $d > 0$ and that the claim holds for $d - 1$. If v is the $\mathbf{f}$-virtual multiplicity of c as a root of $\mathrm{D}_d g$, we have $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(\mathrm{D}_d g); a, b) = v$ by the inductive assumption. We check the different cases, using the definition of $\mathbf{f}$-virtual multiplicity by induction on the $\mathbf{f}$-degree.

- If $g(x)\mathrm{D}_d g(x)$ is negative for $a \leqslant x < c$ and positive for $c < x \leqslant b$, we have $w = v + 1$ and $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b) = \mathrm{V}(\mathrm{Der}_{\mathbf{f}}(\mathrm{D}_d g); a, b) + 1$.
- If $g(x)\mathrm{D}_d g(x)$ does not change sign at c, we have $w = v$ and $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b) = \mathrm{V}(\mathrm{Der}_{\mathbf{f}}(\mathrm{D}_d g); a, b)$.
- If $g(x)\mathrm{D}_d g(x)$ is positive for $a \leqslant x < c$ and negative for $c < x \leqslant b$, we have $w = v - 1$ and $\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b) = \mathrm{V}(\mathrm{Der}_{\mathbf{f}}(\mathrm{D}_d g); a, b) - 1$.

The claim follows in each of these cases. □

Theorem 2 follows from the preceding lemmas and from the obvious additive property: for every $c \in (a, b)$,

$$v(g; (a, b]) = v(g; (a, c]) + v(g; (c, b]),$$
$$\mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, b) = \mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); a, c) + \mathrm{V}(\mathrm{Der}_{\mathbf{f}}(g); c, b).$$

Acknowledgments

We would like to thank Ricky Pollack for his encouragements to our work and his interest in the topic, and Fanny Delebecque and Jimmy Lambolley for useful discussions about the proof. M. Coste, H. Lombardi and M-F. Roy are members of the European RTNetwork Real Algebraic and Analytic Geometry, Contract No. HPRN-CT-2001-00271.

References

[2] R. Descartes, Géométrie (1636). A Source Book in Mathematics, vol. 90-31, Harvard University Press, 1969.
[3] F. Budan de Boislaurent, Nouvelle méthode pour la résolution des équations numériques d'un degré quelconque, (1807), second ed., Paris, 1822.
[4] J. Fourier, Analyse des équations déterminées, F. Didot, Paris, 1831.
[5] L. González-Vega, H. Lombardi, L. Mahé, Virtual roots of real polynomials, J. Pure Appl. Algebra 124 (1998) 147–166.
[6] C. Sturm, Mémoire sur la résolution des équations numériques, u á l'Académie des Sciences le 23 mai 1829.